Deep Koopman

Partie 2 : Architecture + Loss

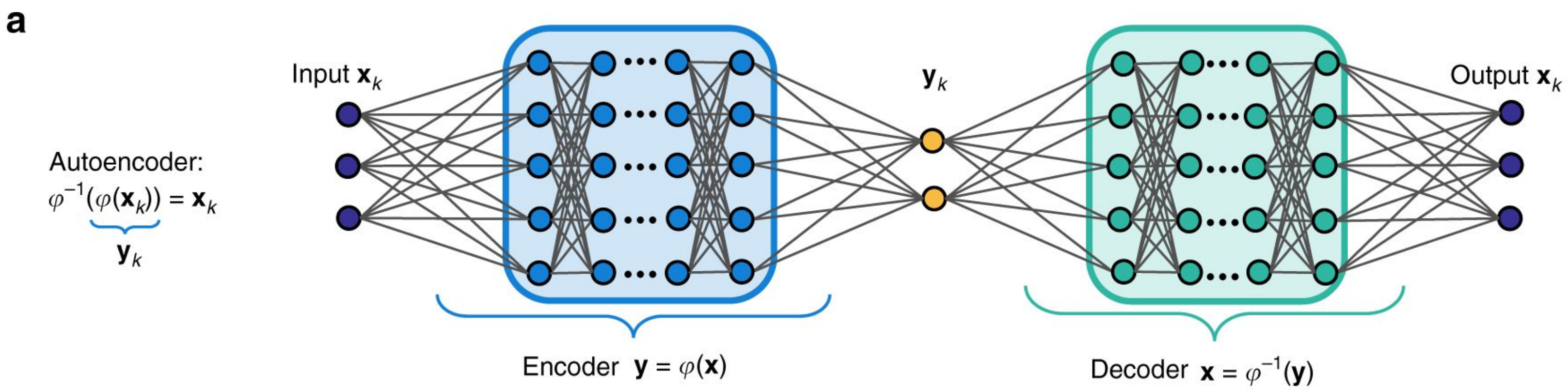
Architecture

Architectures des réseaux de Koopman

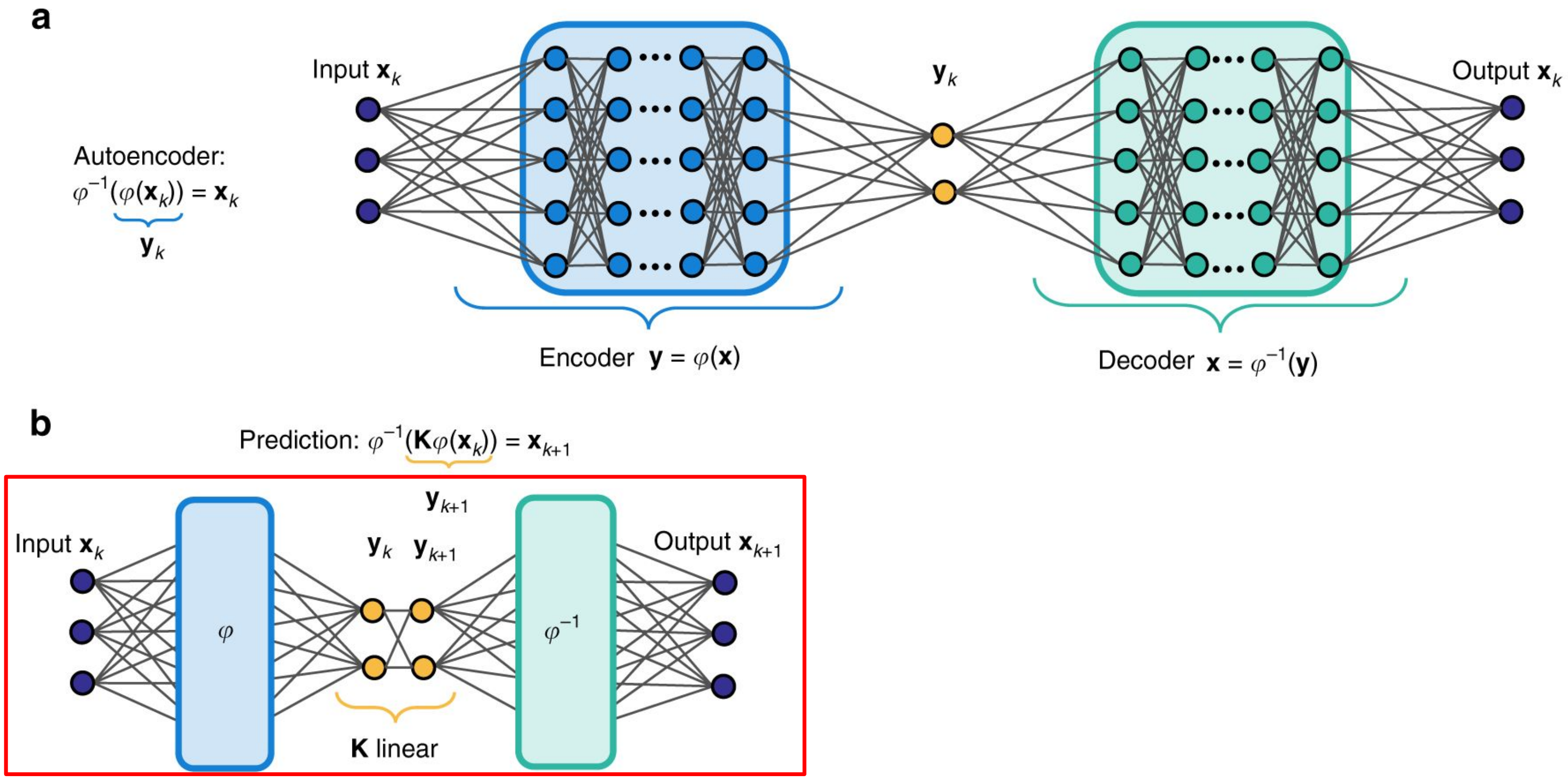
Objectifs :

- Apprendre les fonctions propres de Koopman
    - I.e apprendre espace latent ;
    - Apprendre des fonctions propres inversibles.
- Apprendre les valeurs propres de l’opérateur
    - Ou apprendre la matrice de l’opérateur

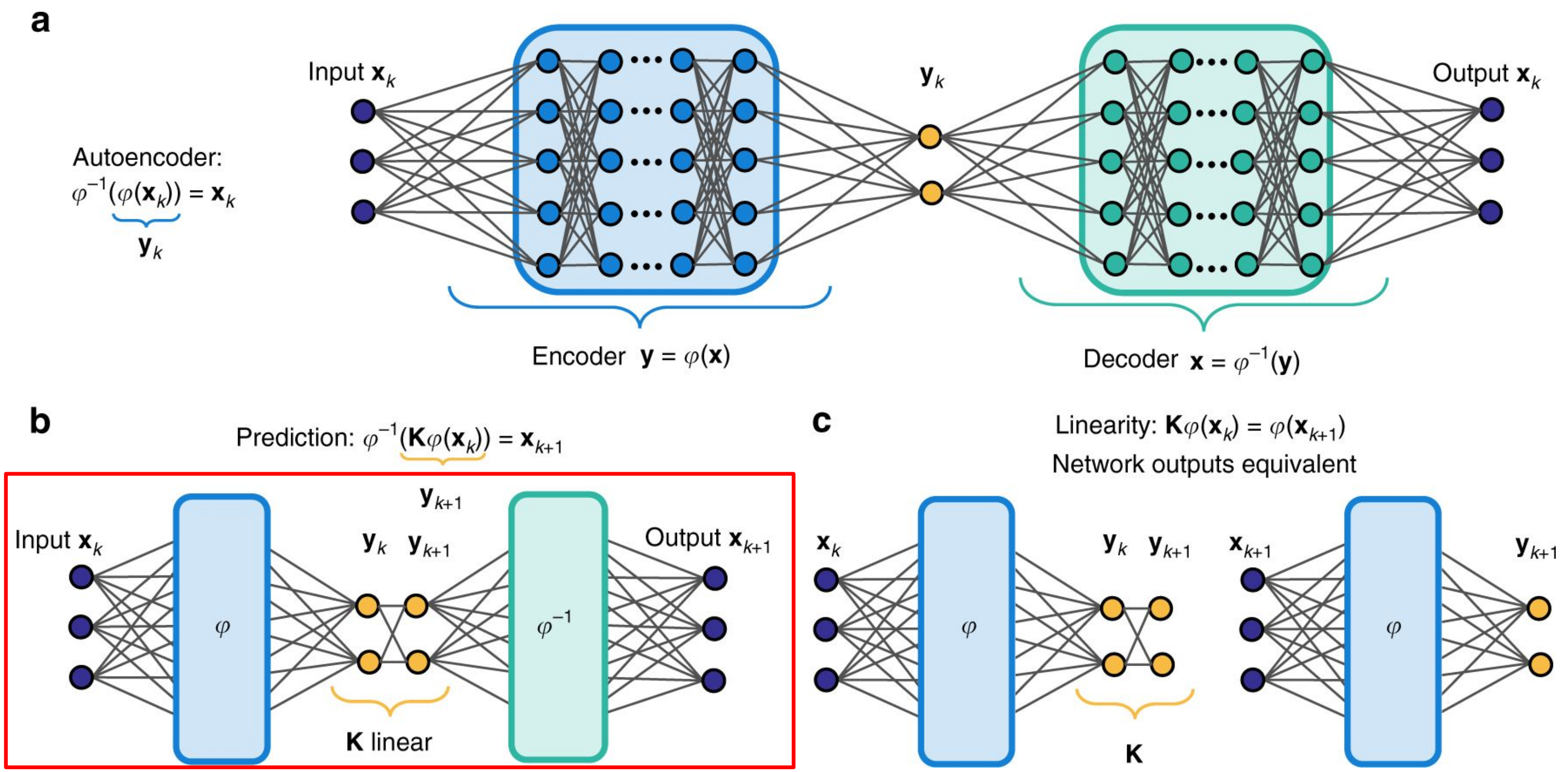
Simple Koopman Autoencoder



Simple Koopman Autoencoder



Simple Koopman Autoencoder



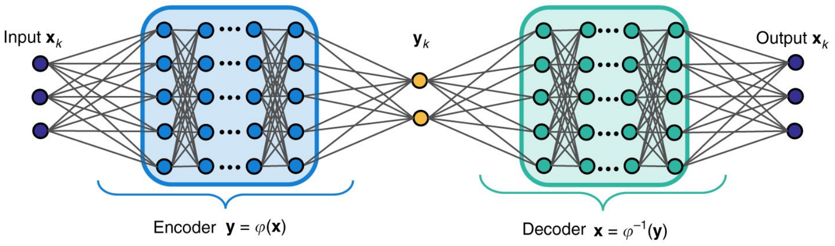
Fonctions Loss

Quelles fonctions de pertes optimiser ?

Objectifs :

- Structure d’Auto-encodeur
- Objectif de prédiction
- Objectif de linéarité

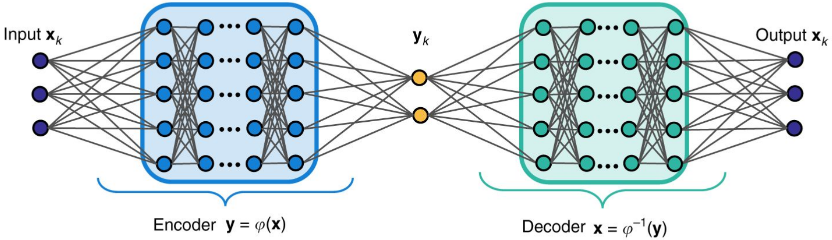
Structure d'Auto-Encodeur



Loss de reconnaissance :

- Mesurer l'écart entre l'input x_k et l'autoencoding en sortie du réseau $\varphi^{-1}(\varphi(x_k))$
- Loss utilisée : MSE

Structure d'Auto-Encodeur

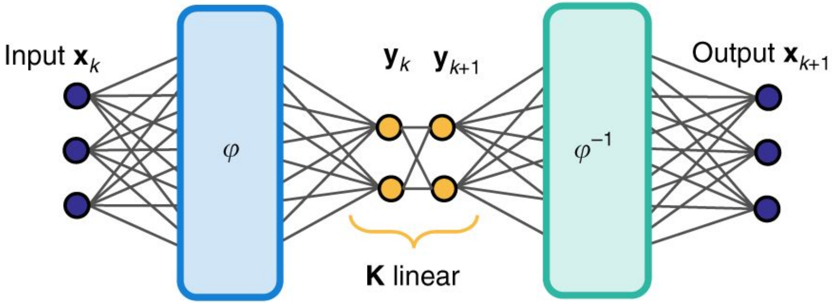


Loss de reconnaissance :

- Mesurer l'écart entre l'input x_k et l'autoencoding en sortie du réseau $\varphi^{-1}(\varphi(x_k))$
- Loss utilisée : MSE

$$\mathcal{L}_{\text{recon}} = \left\| \mathbf{x}_1 - \varphi^{-1}(\varphi(\mathbf{x}_1)) \right\|_{\text{MSE}}$$

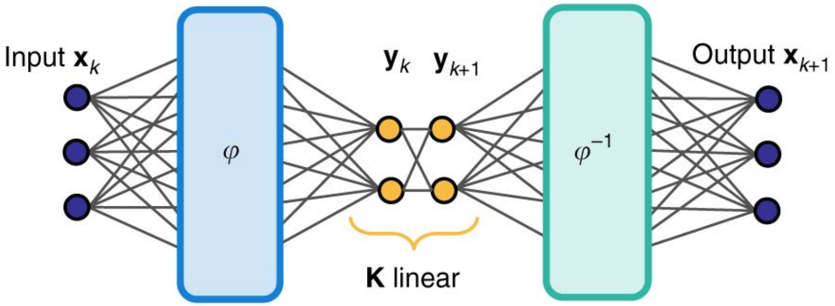
Prédiction



Loss de prédiction :

- Objectif : Prévoir l'instant suivant x_{k+1} à partir d'un instant x_k

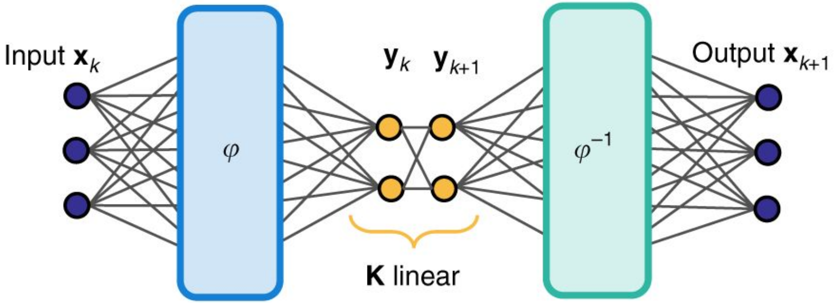
Prédiction



Loss de prédiction :

- Objectif : Prévoir l'instant suivant x_{k+1} à partir d'un instant x_k
  - Loss utilisée : MSE
  - Loss = $||x_{k+1} - \varphi^{-1}(\mathbf{K}\varphi(x_k))||_{MSE}$

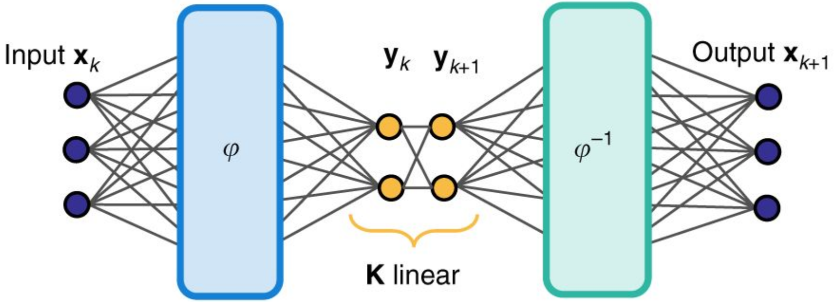
Prédiction



Loss de prédiction :

- Objectif : Prévoir l'instant suivant x_{k+1} à partir d'un instant x_k
  - Loss utilisée : MSE
  - Loss = $||x_{k+1} - \varphi^{-1}(\mathbf{K}\varphi(x_k))||_{MSE}$
- Amélioration : Prévoir une trajectoire de plusieurs pas à partir d'un instant initial x_1

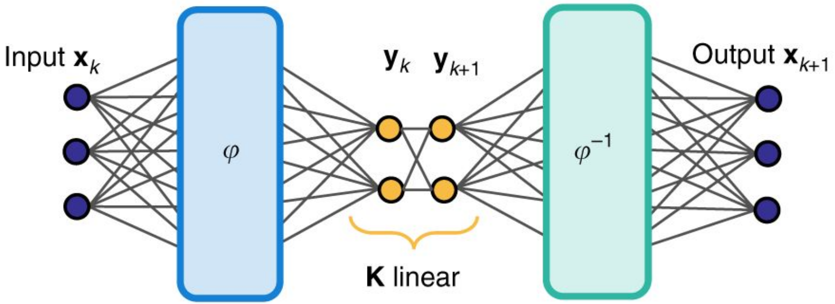
Prédiction



Loss de prédiction :

- Objectif : Prévoir l'instant suivant x_{k+1} à partir d'un instant x_k
  - Loss utilisée : MSE
  - Loss = $||x_{k+1} - \varphi^{-1}(\mathbf{K}\varphi(x_k))||_{MSE}$
- Amélioration : Prévoir une trajectoire de plusieurs pas à partir d'un instant initial x_1
  - On a la relation : $x_{m+1}=\varphi^{-1}(\mathbf{K}^m \varphi(x_1))$
  - Hyper paramètre supplémentaire : Longueur de la trajectoire prédite S_P

Prédiction

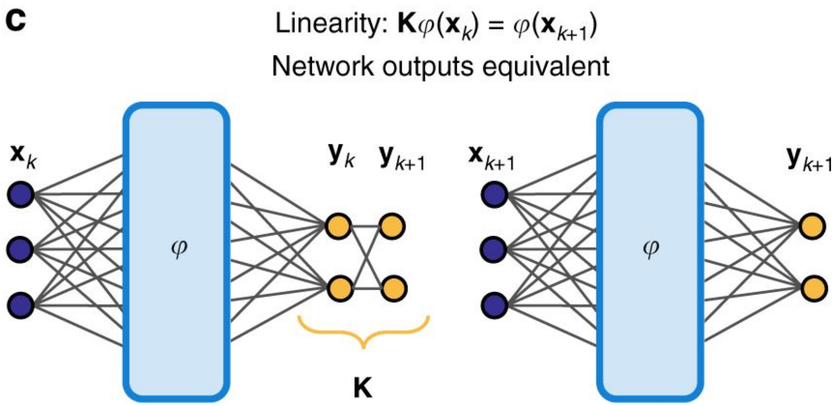


Loss de prédiction :

- Objectif : Prévoir l'instant suivant x_{k+1} à partir d'un instant x_k
  - Loss utilisée : MSE
  - Loss = $||x_{k+1} - \varphi^{-1}(\mathbf{K}\varphi(x_k))||_{MSE}$
- Amélioration : Prévoir une trajectoire de plusieurs pas à partir d'un instant initial x_1
  - On a la relation : $x_{m+1}=\varphi^{-1}(\mathbf{K}^m\ \varphi(x_1))$
  - Hyper paramètre supplémentaire : Longueur de la trajectoire prédite S_P
- Loss Finale :

$$\mathcal{L}_{\text{pred}} = \frac{1}{S_p}\sum_{m=1}^{S_p}\left\|\mathbf{x}_{m+1} - \varphi^{-1}(K^m\varphi(\mathbf{x}_1))\right\|_{\text{MSE}}$$

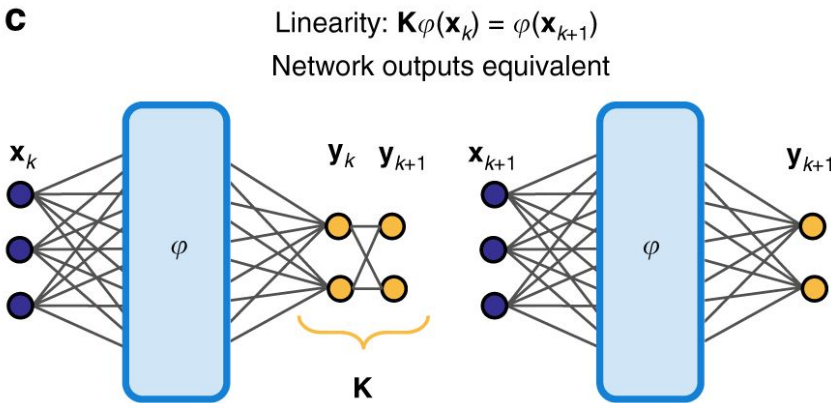
Linéarité



Loss de linéarité :

- Objectif : apprendre l'opérateur de Koopman **K**

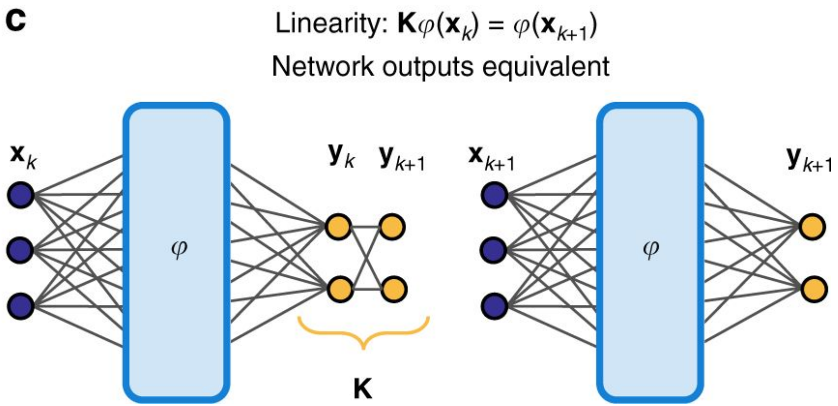
Linéarité



Loss de linéarité :

- Objectif : apprendre l'opérateur de Koopman **K**
  - Loss utilisée : MSE
  - Loss = $||\varphi(x_{k+1}) - \mathbf{K}\varphi(x_k)||_{MSE}$

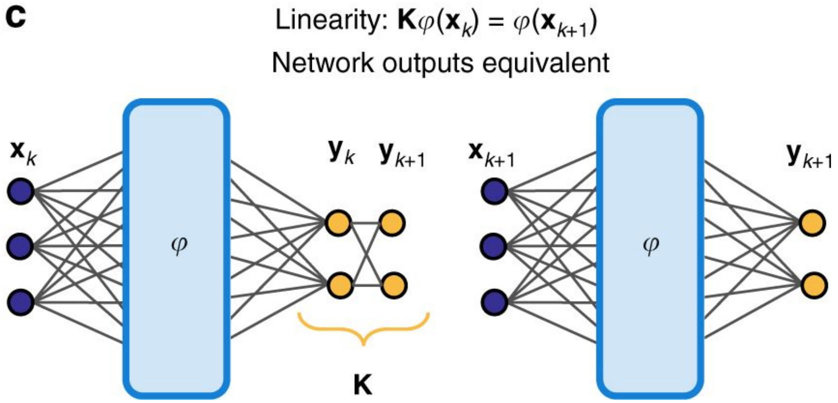
Linéarité



Loss de linéarité :

- Objectif : apprendre l'opérateur de Koopman **K**
  - Loss utilisée : MSE
  - Loss = $||\varphi(x_{k+1}) - \mathbf{K}\varphi(x_k)||_{MSE}$
- Amélioration : Prédiction linéaire sur toute la trajectoire

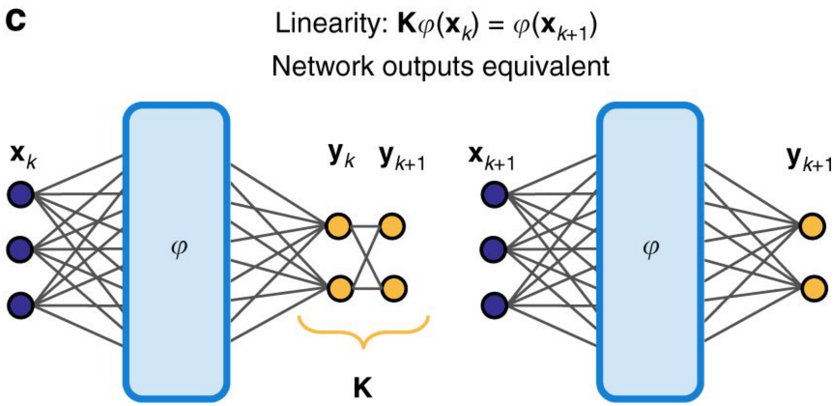
Linéarité



Loss de linéarité :

- Objectif : apprendre l'opérateur de Koopman **K**
  - Loss utilisée : MSE
  - Loss = $||\varphi(x_{k+1}) - \mathbf{K}\varphi(x_k)||_{MSE}$
- Amélioration : Prédiction linéaire sur toute la trajectoire
  - Relation $\varphi(x_{m+1}) = \mathbf{K}^m\varphi(x_1)$
  - Longueur de la trajectoire T (fixée)

Linéarité



Loss de linéarité :

- Objectif : apprendre l'opérateur de Koopman **K**
  - Loss utilisée : MSE
  - Loss = $||\varphi(x_{k+1}) - \mathbf{K}\varphi(x_k)||_{MSE}$
- Amélioration : Prédiction linéaire sur toute la trajectoire
  - Relation $\varphi(x_{m+1}) = \mathbf{K}^m\varphi(x_1)$
  - Longueur de la trajectoire T (fixée)
- Loss finale :

$$\mathcal{L}_{\text{lin}} = \frac{1}{T-1}\sum_{m=1}^{T-1} \|\varphi(\mathbf{x}_{m+1}) - K^m\varphi(\mathbf{x}_1)\|_{\text{MSE}}$$

Loss additionnelles

- Pas d'outliers dans les données
- Ajout d'une loss infinie pour pénaliser le plus grand écart sur les loss de reconnaissances et prédictions

$$\mathcal{L}_\infty = \left\|\mathbf{x}_1 - \varphi^{-1}(\varphi(\mathbf{x}_1))\right\|_\infty + \left\|\mathbf{x}_2 - \varphi^{-1}(K\varphi(\mathbf{x}_1))\right\|_\infty$$

- Pour éviter l'overfitting : Ajout d'une Loss l^2 sur les poids du modèle $\|\mathbf{W}\|_2^2$

Loss additionnelles

- Pas d'outliers dans les données
- Ajout d'une loss infinie pour pénaliser le plus grand écart sur les loss de reconnaissances et prédictions

$$\mathcal{L}_\infty = \left\|\mathbf{x}_1 - \varphi^{-1}(\varphi(\mathbf{x}_1))\right\|_\infty + \left\|\mathbf{x}_2 - \varphi^{-1}(K\varphi(\mathbf{x}_1))\right\|_\infty$$

- Pour éviter l'overfitting : Ajout d'une Loss l^2 sur les poids du modèle $\|\mathbf{W}\|_2^2$

Loss finale : Combinaison linéaire des Loss précédentes

- 3 Hyperparamètres supplémentaires

$$\mathcal{L} = \alpha_1(\mathcal{L}_{\text{recon}} + \mathcal{L}_{\text{pred}}) + \mathcal{L}_{\text{lin}} + \alpha_2\mathcal{L}_\infty + \alpha_3\|\mathbf{W}\|_2^2$$

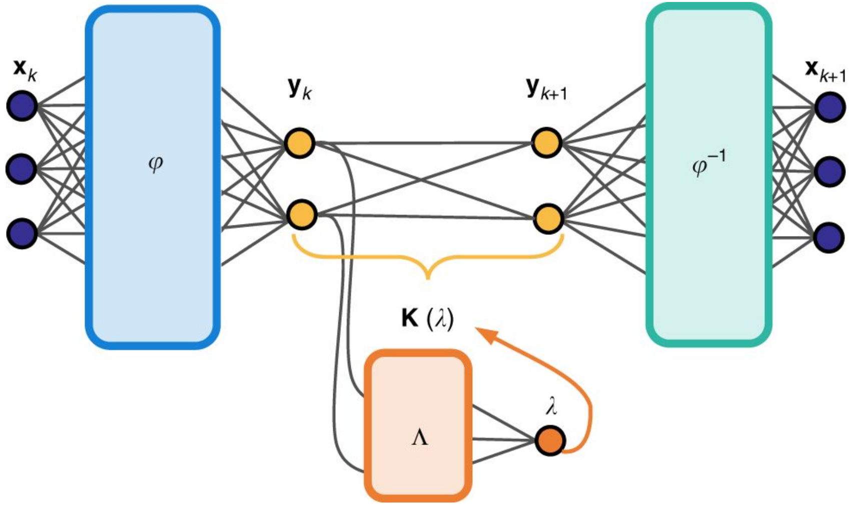
Variation de l'autoencodeur de Koopman

Une méthode plus générale

Objectifs :

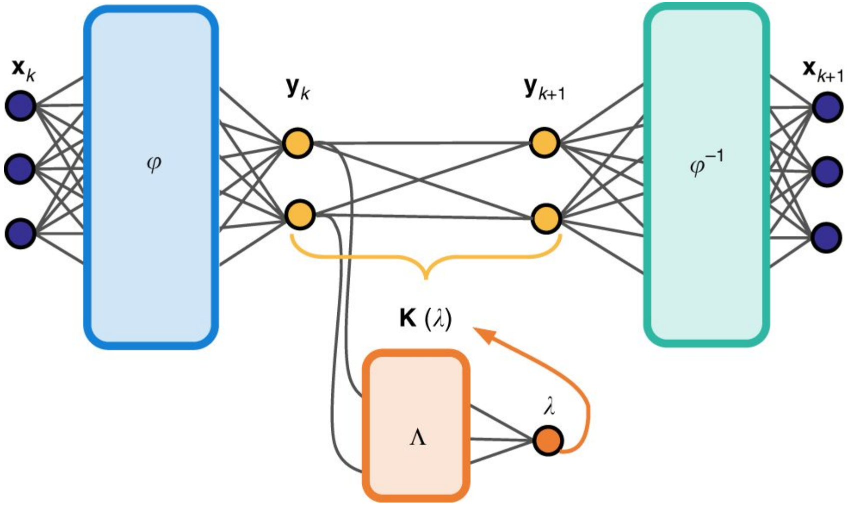
- Cas où le spectre de valeurs propres est continu
  - Comment représenter la matrice de Koopman ?

Auto Encodeur de Koopman Complet



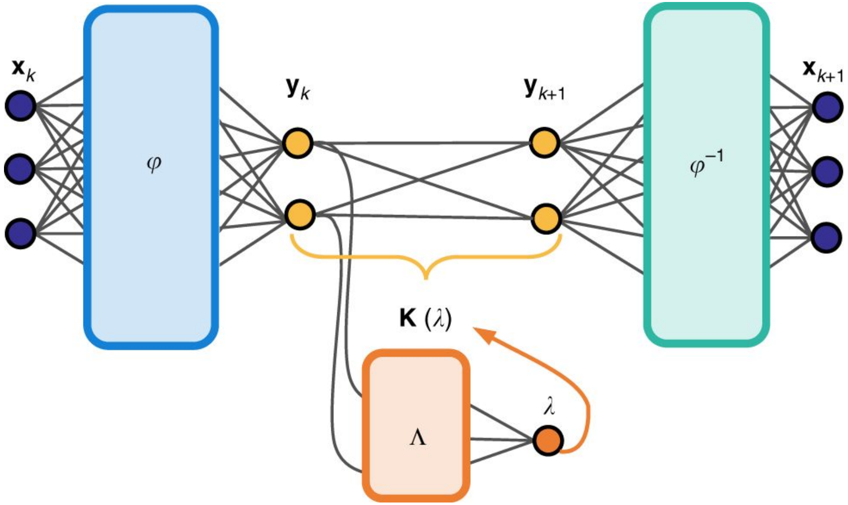
- Apprentissage directe des valeurs propres de l’opérateur **K** par un réseau auxiliaire
  - Réinjection dans le réseau directement

Auto Encodeur de Koopman Complet



- Apprentissage directe des valeurs propres de l’opérateur **K** par un réseau auxiliaire
  - Réinjection dans le réseau directement
- Avantages :
  - Permet de réduire la dimension de l’espace latent ;
  - Permet de considérer le cas où le spectre de valeurs propres est continu ;
  - Englobe le cas simple (Pas de perte de généralisation)

Auto Encodeur de Koopman Complet



- Apprentissage directe des valeurs propres de l’opérateur **K** par un réseau auxiliaire
  - Réinjection dans le réseau directement
- Avantages :
  - Permet de réduire la dimension de l’espace latent ;
  - Permet de considérer le cas où le spectre de valeurs propres est continu ;
  - Englobe le cas simple (Pas de perte de généralisation)
- Structure et Loss ?
  - Paramétrisation de K par un réseau MLP auxiliaire
  - Pas de changements de Loss